INFO 290T: Written Report

An exploration of biological response to anthropomorphized stimuli using Gaze-Tracking and EDA biofeedback.

Paul Kenniston, Erin Engle Jones & Shikha Shah
University of California, Berkeley

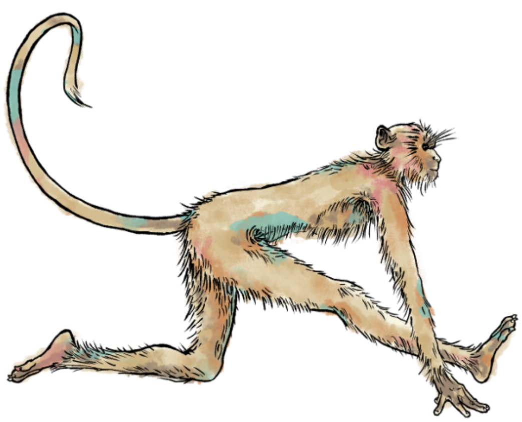


Table of Contents

Abstract

We used three different stimulus types (each corresponding to a different expression of anthropomorphization) to record participant biosignal responses in order to determine whether responses differ systematically across stimulus groupings. If we identify a significant difference in participant response across these stimulus groups, we can make inferences about how people respond to anthropogenic stimuli more generally.

Our study on the impact of anthropomorphization in visual and stimuli, using biosensory analysis of pupil data and EDA responses, reveals that elements like backstory significantly enhance recall, with a notable participant remembering specific dialogues. However, challenges such as participant fatigue and data normalization issues were encountered. The study suggests the potential for applying these findings in attention-dependent fields like pedagogy and raises intriguing questions about the broader implications of anthropomorphization in cognition and human perception.

The three different groups of anthropogenic stimuli are listed below:*

Type 1 [Control]: inanimate objects with faces,
Type 2 [Moniker]: inanimate objects with names/greetings provided via text,
Type 3 [Backstory]: inanimate objects with AI-generated audio greetings

** See 'Appendix A' for our collection of stimulus images, monikers, and associated stories.*

Keywords

Anthropomorphism; pedagogy; biosensory; electrodermal activity; eye-tracking; attention; engagement

Introduction

Anthropomorphization and teleological explanations for objects and processes have been hotly debated in pedagogical research, with some factions asserting that the use of such tactics simply confuses young students (Kallery et al), while others assert that use of these techniques may actually help with understanding and recall, especially in more mature populations (Stoos et al).

Our aim is to investigate the variance in biosignal response as it relates to different types of anthropomorphization. By measuring biofeedback, we can quantify our participants' responses to various forms of anthropomorphic stimuli. If we assume that there is an intrinsic relationship between biosensory activity and ‘engagement’ or ‘attention’, we can then extrapolate our results and generate inferences about which forms of anthropomorphism garners the most attention.

These inferences can then be applied in industries that are dependent on ‘attention’. One field that is dependent on attention that could utilize this information is that of pedagogy (Chiong et al, Larsen et al). Teachers could utilize anthropomorphic stimuli in their lesson plans in order to keep their student’s attention during long lessons. Or, teachers could associate difficult to understand or remember concepts with a story in order make the concept more relatable to their students.

The findings may also be applicable to discourse surrounding cognition and dehumanization (Waytz et al). When we anthropomorphize our surroundings, we are attributing both physical and non-physical human characteristics to the subject. We might associate the subject with a human face, but we may also attribute emotions (e.g., joy, grief, sadness), cognitive awareness, or moral intent to the subject. There are wide ranging consequences and implications of non-human subjects being perceived as having complex mental faculties. Non-human subjects “that are capable of judgment, intention, and feeling are also capable of directing their judgment, intentions, and feelings toward us, and therefore become agents of social influence” (Waytz et al). Further, if we collectively perceive that a non-human subject is capable of complex mental faculties, does that make the subject worthy of moral care and protection that are typically reserved for humans (Waytz et al)? While our experiment evaluating participants’ responses to anthropomorphization has direct utility in attention-dependent fields like pedagogy, it also has the potential to speak to larger concerns regarding the human condition and possible extensions of what it means to be human.

Methods

Experimental Methodology

To conduct this study, we presented our participants with three types of stimuli in blocks, as shown in Appendix A, and summarized in the following paragraph. Participants will be wearing eye-tracking (*Pupil Invisible)* and EDA (*Empatica Embrace+*) sensors during the experiment in order for us to receive a variety of possible biosensor data that we can analyze against the hypotheses stated above. Each sensor was calibrated prior to gathering data to account for variations in the participant’s distance to the display showing the stimuli video, lighting in the room, and other fluctuations caused by the hardware. Our methodology for calibration is given in the ‘*Calibration Methodology*’ section below.

We will present our participants with 3 different categories of anthropomorphized images, each containing the same 15 images, but with slight modifications across groups. Each image (a unit in a block of stimuli) will be shown for 10 seconds. Each image shows a non-human object or scene, but contains something that resembles eyes (to create a focal point for fixation tracking).

These images have been rendered as six videos, each representing one of different permutations of group types, each video and recording session lasting approximately 7 minutes.

One stimuli category represents a control group of ‘base’ anthropomorphic characteristics, solely containing images without any additional context. This is our control group. Our ‘named’ group is composed of images presented alongside a moniker and a greeting in written form. The final group, our ‘backstory’ group, contains images presented with an audio track that creates a backstory for the object in the image. To account for the order in which images are presented, we will randomize (with removal) the order that each category is presented to participants.

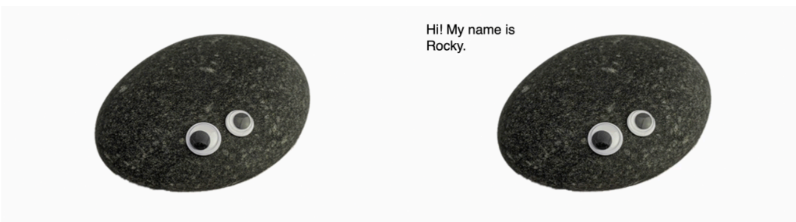


Figure 1: Control image (left) vs image with written name (right)

To collect data, the recorder will place both sensors on the participant. Once the devices are on, the recorder will begin the recording on *Empatica Embrace+* while simultaneously starting a stopwatch. At the 15 second mark, we begin recording the Pupil Invisible. At the 30 second mark, we begin the stimulus video. The video shown will correspond with a random number generated between 1 and 6 (inclusive) by the recorder, which randomizes the order in which the participant sees each group of stimuli. At the end of the stimulus video, as soon as the white slide shows, we wait an additional 15 seconds then end the recording on the *Pupil Invisible*. Finally, after another 15 seconds passes (30 seconds after the video ends), the recorder will stop the recording on the *Empatica Embrace+*. Should any of these devices not work during the recording, it shall be marked in the participant log and device start/end recording shall still be spaced by 15 second increments.

After collecting data, we conducted analyses of both sets of data from all the three devices across participants using the methods applied in Assignments 1, 2 and 4 to see if participant reactions differed enough to merit the rejection of any of our null hypotheses. We conducted this data analysis exclusively in Python, and we utilized existing packages, as documented in the ‘Materials’ section.

During our data analysis, we sought to identify significant differences in the summary statistics between the various stimuli categories.Our analysis methods generated a set of summary features for each image viewed. Our findings on participant's responses to these stimuli sets and our summary of the importance of specific categories of biosensory data is documented in the 'Results' section of this report.

Calibration Methodology

Methods for calibration and standardization when using the Pupil Invisible eye tracker:

- Perform camera calibration - This involves having the user look at QR code dots in predefined locations to map their gaze to the camera footage. The Pupil software has a built-in calibration routine.
- Check headband fit - The headband needs to be fitted securely and the eye cameras positioned correctly in front of the eyes. This affects calibration accuracy.
- Set viewing distance and angle - Standardize the distance and angle between the user and screen/object of regard. This affects the mapping between gaze and tracking area.
- Stabilize head position - Use a chin rest to minimize head movements after calibration. Major head movements can interfere with the original gaze mapping calibration.

Methods for calibration and standardization when collecting EDA (electrodermal activity) data with the Empatica E4 wristband:

- Sensor placement - The wristband needs to be worn on the non-dominant hand, fitted snugly, at least 1 finger-width above the wrist bone, with the sensors on the inside of the wrist making good contact with the skin. Consistent placement is key.
- Hydration - Ensure participants are well hydrated, as dehydration can affect EDA signals. Avoid alcohol, caffeine or vigorous exercise immediately before recording.
- Participant guidelines - Instruct participants to limit hand movements and stay seated during baseline and recording. Movement can create signal artifacts.

Materials

Below is a comprehensive list of materials used during this experiment.

- Images (sourced ourselves from stock images, Google)
- AI generated audio (generated using free trial on Design AI)
- Laptop with the ability to connect to selected devices
- Eye-tracking sensor [*Pupil Invisible*]
- EDA sensor [*Empatica Embrace+*]
- 6(+) participants
- Python libraries for data analysis
    - *pandas*
    - *numpy*
    - *H5py*
    - *Scipy.signal*
    - *Neurokit2*

- *Plotly*
- *Matplotlib.pyplot*
- *Seaborn*
- *avro*

Challenges

We hit significant hardware issues and several methodological issues after recording our first participant. We ended up tossing the data from that recording session, but documented several challenges and lessons that defined our procedure moving forward. These roadblocks and challenges are summarized below and led to the following changes in methodology and equipment as summarized below:

Challenge	**Solution/Modification**
Participant 1, recorded 11/2, experienced significant fatigue by the end of the video as we initially were showing each image for 20 seconds instead of 10 seconds.	The stimulus video was updated to show each image for only 10 seconds.
Participant 1's pupil tracking data was largely unusable, due to calibration issues.	Calibration procedures and standardization methodology are now included in methods to hopefully produce more usable data
Shikha and Erin were unable to connect the Neurosky at all in the initial experimental design/solidification meeting on 11/1 as neither had a Window's machine.	During our meeting on 11/7, we were able to connect Erin's device to the Neurosky and as such will work to collect this type of data. As a safety net, we also checked out the Empatica Embrace+, which produces reliable data per Erin's hob experience with the device. This will allow Shikha to record in the event that her device will still not connect to the Neurosky and will offer us some flexibility in data collected in the event that any of the data is less than usable.
Shikha and Erin both sourced windows machines but still could not connect the neurosky after significant troubleshooting methods (changing batteries, spinning up new admin instances, etc) - they could install the software but the device would not connect via bluetooth. It would pair but immediately drop the connection.	See response above. This was a setback in both the number of individuals we were able to record and our timeline, but we do not feel that it prevents us from completing our experiment on time.
"Anthropomorphism" can be expressed in a variety of different ways, and 'images resembling faces' is merely one possible	During early discussions in which we were defining methodology, we had different interpretations of how to best depict

interpretation of a much grander concept. The ways that we chose to represent anthropomorphism (e.g. ‘named images’, ‘backstories presented via audio’) may not capture the bigger picture. Further, the results of our experiment may not be applicable to the entirety of anthropomorphic-driven educational practices.	anthropomorphism. We debated whether audio was relevant, whether depictions of human-made items (e.g. a mask) qualify as being anthropomorphic, and whether anthropomorphism requires some sort of active human intervention (e.g. putting eyes on a rock). We ended up altering our original research question to indicate that our results only speak to a narrow slice of anthropomorphism, as we realized that creating an experiment that could capture the depth of the concept was beyond the possible scope of this project.
Pupil Invisible turned out to be far more light-sensitive than we had originally anticipated. Several of our recording sessions were conducted in seemingly well-lit rooms with lights fixated on a high ceiling. The light-level in this room was not enough for the PupilLabs hardware to consistently recognize the QR codes adjacent to the stimuli. As a result, the data that we received for these participants was not ‘grounded’ to the screen that was presenting the stimuli.	The data that we received in these conditions prevented us from knowing where exactly participants were looking at (ideally, we would know what parts of the stimuli kept their gaze). However, we were still able to get absolute measures of saccades and fixations, regardless of where those saccades and fixations took place. We used these screen-ambivalent results in our data analysis, rather than the screen-relative data we were originally hoping to get.

Results

Commented [1]: Fill in after analysis

Quantitative Results: Eye-tracking

Using the Pupil Invisible, we were able to gather data on the quantity and length of gaze fixations and saccades of our participants when viewing the stimuli. We arrived at the following metrics from the raw data, which we believe will allow us to effectively evaluate the validity of our hypothesis:

For analyzing fixations:

- 'Total fixations'
- 'Average fixation length'*
- 'Total long fixations'**
- 'Average long fixation length'

For analyzing saccades:

- 'Total saccades'
- 'Total horizontal saccades'
- 'Total vertical saccades'
- 'Ratio of horizontal to vertical saccades'

** We set our Pupil Invisible to capture any fixation greater than 60ms.*

** *A "long" fixation is equal to or greater in length than two standard deviations above the participant's average fixation length.*

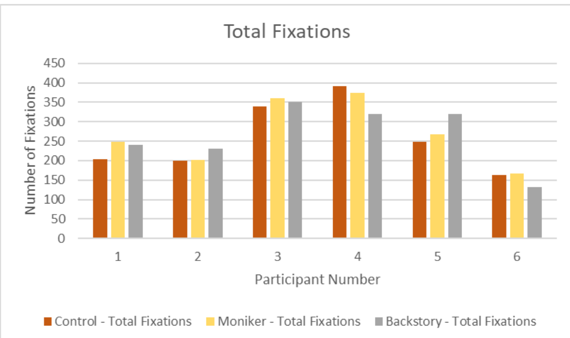
As mentioned above in our 'Challenges' section, we decided to rely on these absolute metrics (rather than on-screen location of the participant's gaze) due to quality control issues in our recording environment. Due to the hardware's sensitivity to lighting, we were only able to get relative screen-specific data for 2 of our 7 participants. Additionally, we had to throw out one of our recordings due a hardware error that occurred midway through a session.

Below are the processed Pupil Invisible results of our 6 participants:

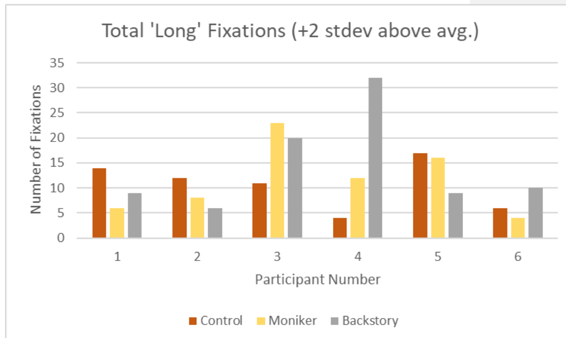
Participant Number	1	2	3	4	5	6
Fixation Analysis:						
Control - Total Fixations	204	199	339	391	248	163
Control - Avg. Fixation Length (ms)	637.8578	715.9447	295.7463	257.2864	499.9637	804.7975
Control - Total Long Fixations	14	12	11	4	17	6
Control - Avg. Long Fixation Length (ms)	2227.643	3309.917	1252.273	1004.75	1972.765	3396.167
Moniker - Total Fixations	249	201	360	374	267	166
Moniker - Avg. Fixation Length (ms)	517.6265	692.0647	332.6111	310.6818	464.4382	798.0542
Moniker - Total Long Fixations	6	8	23	12	16	4
Moniker - Avg. Long Fixation Length (ms)	2386.667	3875.125	1178.478	1193.333	1610.625	3555.5
Backstory - Total Fixations	240	230	350	319	320	133
Backstory - Avg. Fixation Length (ms)	526.1583	607.9087	352.4057	425.4796	379.8906	1040.03
Backstory - Total Long Fixations	9	6	20	32	9	10
Backstory - Avg. Long Fixation Length (ms)	1867.222	2673.333	1130.35	1235.969	1431.889	4713.6
Saccade Analysis						
Control - Total Saccades	203	198	338	390	247	162
Control - Total Horizontal Saccades	93	119	237	204	153	100
Control - Total Vertical Saccades	110	79	101	186	94	62
Control - Horizontal to Vertical Ratio	0.8455	1.5063	2.3465	1.0968	1.6277	1.6129
Moniker - Total Saccades	248	200	359	373	266	165
Moniker - Total Horizontal Saccades	156	127	267	209	169	104
Moniker - Total Vertical Saccades	92	73	92	164	97	61
Moniker - Total Vertical Saccades	92	73	92	164	97	61
Moniker - Horizontal to Vertical Ratio	1.6957	1.7397	2.9022	1.2744	1.7423	1.7049
Backstory - Total Saccades	239	229	349	318	319	132
Backstory - Total Horizontal Saccades	127	106	241	172	201	85
Backstory - Total Vertical Saccades	112	123	108	146	118	47
Backstory - Horizontal to Vertical Ratio	1.1339	0.8618	2.2315	1.1781	1.7034	1.8085

Figure 2: Compilation of eye-tracking results

Using this raw data, we were able to create some visualizations that better represent how our participant's biodata differed across categories. Below is a collection of charts that visualize the data presented in the table above:



Total Fixations
Number of Fixations
450
400
350
300
250
200
150
100
50
0
1
2
3
4
5
6
Participant Number
Control - Total Fixations
Moniker - Total Fixations
Backstory - Total Fixations



Total 'Long' Fixations (+2 stdev above avg.)
Number of Fixations
35
30
25
20
15
10
5
0
1
2
3
4
5
6
Participant Number
Control
Moniker
Backstory

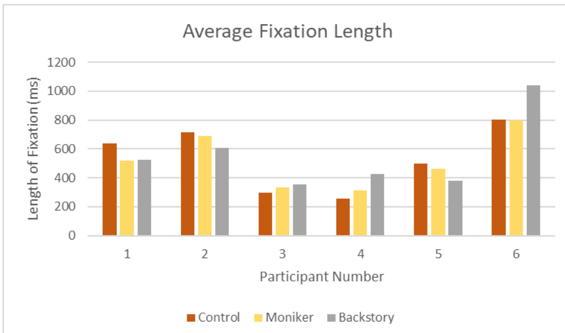


Figure 3: Visualizations of Fixation Data Captured from Pupil Invisible

Top left: 'Total Fixations'; Top Right: 'Total Long Fixations'; Bottom Middle: 'Average Length of Fixations (ms)'

Reviewing the visualization of fixations alone, we noticed a few trends that reflect how participants reacted differently to different stimuli categories.

For most of our participants, the total number of fixations increased in the two experimental groups as compared to the control group. However, this difference is marginal, and for several of our participants there was a decrease in fixations in the experimental groups. The biggest delta between either of the two experimental groups and the control group occurred during Participant 5's recording, who had a 29% increase in fixations in the 'Backstory' group as compared to the control group. Participant 1 also experienced a 22% increase in fixations in the 'Moniker' group as compared to the control group. As for those who experienced a decrease in fixations in the experimental groups - both Participant 4 and Participant 6 experienced a 19% decrease in fixations in the 'Backstory' group. Beyond these few notable exceptions, the difference between control and experimental groups was within the bounds of -5% to +17%.

Our results on 'Average Fixation Length' were a slightly better indicator of there being differences in biofeedback to different anthropomorphized stimuli. While this metric speaks to more extreme differences between control and experimental groups, it shows extreme differences in both the positive and negative directions. Some of our participants had very dramatic increases in the length of their fixations in the experimental groups (such as Participant 4, who had a 65% increase in fixation length in the 'Backstory' category), while others had significant decreases in fixation length (such as Participant 5, who had a 25% decrease in the 'Backstory' category).

Control' : 'Moniker' Comparison		
Ratio of Total Fixations	Ratio of Fixation Length	Ratio of Long Fixations
1.2206	0.8115	0.4286
1.0101	0.9666	0.6667
1.0619	1.1247	2.0909
0.9565	1.2075	3.0000
1.0766	0.9289	0.9412
1.0184	0.9916	0.6667

Figure 4: Eye-tracking data, 'Control' group compared to 'Moniker' group

Control' : 'Backstory' Comparison		
Ratio of Total Fixations	Ratio of Fixation Length	Ratio of Long Fixations
1.1765	0.8249	0.6429
1.1558	0.8491	0.5000
1.0324	1.1916	1.8182
0.8159	1.6537	8.0000
1.2903	0.7598	0.5294
0.8160	1.2923	1.6667

Figure 5: Fixation data, 'Control' group compared to 'Backstory' group

Both the 'Total Fixations' metric and the 'Average Length of Fixation' metric show that participants may react differently to different categories of anthropomorphization; however, reactions seen in these metrics are so inconsistent that it is difficult to parse any greater trend that spans across participants.

However, once we took a step back and considered the differences between fixations rather than flatting all fixations to a single category, we were able to parse some greater nuance from our datasets. We created a new sub-category of fixations called "long" fixations, which includes any fixation that is longer than 2 standard deviations above the participant's average. We believe that this category is a reasonable facsimile of a participant's 'attention' - the assumption behind this is that participants who are more engaged with their stimuli will hold eye contact with the stimulus for longer periods of time.

Distinguishing between fixation types revealed even more drastic differences between our experimental and control groups; but, these differences also seem to be dependent on the person, and like the other metrics, have very inconsistent directionality. For example, Participant 4 had a 200% increase in long fixations in the ‘Moniker’ group and a 700% increase in long fixations in the ‘Backstory’ group. But, Participant 1 had a 58% decrease in long fixations in the ‘Moniker’ group and a 36% decrease in the ‘Backstory’ group.

Our analysis of our participants’ saccades yielded similar results to our fixation analysis. While participants’ eye movements were slightly different in the experimental groups than in the control group, the directionality of that change was inconsistent, and the total amount of saccades was only marginally higher or lower.

Control : Moniker Comparison
Ratio of Total Saccades
0.818548387
0.99
0.941504178
1.045576408
0.928571429
0.981818182

Control : Backstory Comparison
Ratio of Total Saccades
0.849372385
0.864628821
0.968481375
1.226415094
0.774294671
1.227272727

Figure 6: Saccade data

Left: ‘Control’ group compared to ‘Moniker’ group
Right: ‘Control’ group compared to ‘Backstory’ group

Our eye-tracking results show that our participants do fixate in different ways to different types of anthropogenic stimuli. But, given our small sample size, and the differences in directionality of any changes that occurred between our control and experimental groups, we can’t say whether these changes in fixation habits are purely individualistic or if they speak to a greater trend that is applicable to humans at large.

Quantitative Results: EDA

EDA was captured using the Empatica EmbracePlus on the user’s non-dominant hand. User’s hands rested on the table in front of them to prevent any noisy data from interfering with the results collected. This device also recorded temperature and blood volume pulse, however, our team opted to focus on the EDA in hopes of capturing differing degrees of reaction to the various stimuli.

The results for the EDA were limited to 6 participants, as the 7th participant had an issue where the device continued recording and due to the imprecise nature of the estimated start time for that individual, we were unable to determine when the start time for the stimuli was given the timed nature of the results, so this user was abandoned.

The below three plots show in Figure 7 the distribution of EDA mean and median values (normalized) for each period by the stimuli group they were shown. Our hope was to see an overall lower normalized EDA for the control group than for both the moniker and backstory. This does seem to be the case for the first participant and the fourth participant, however the general shape of the graphs remains the same and the scale does not shift significantly.

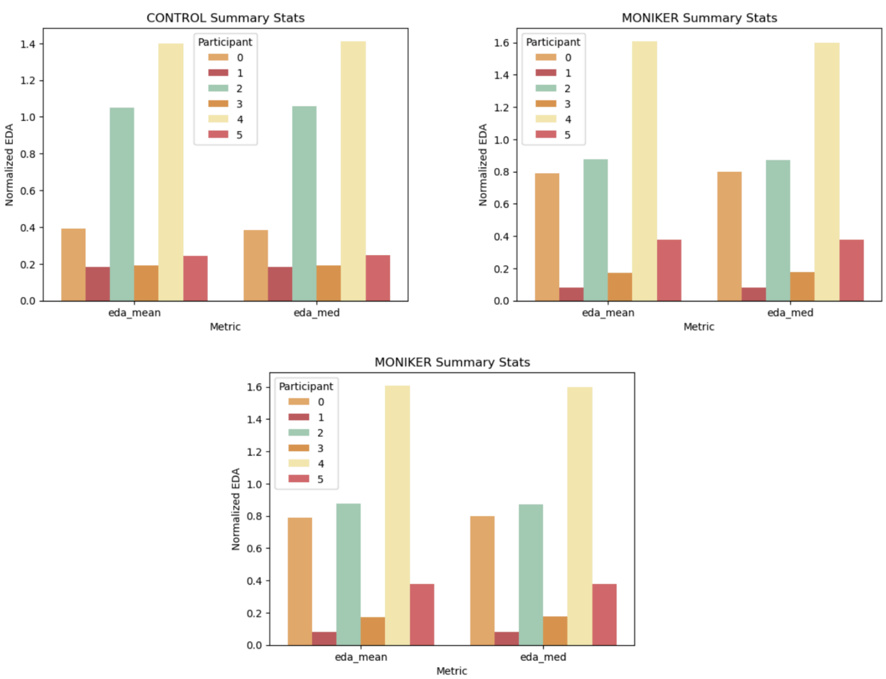
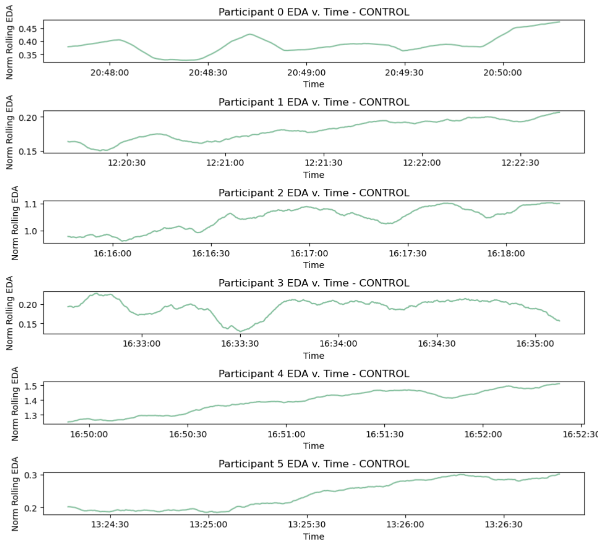
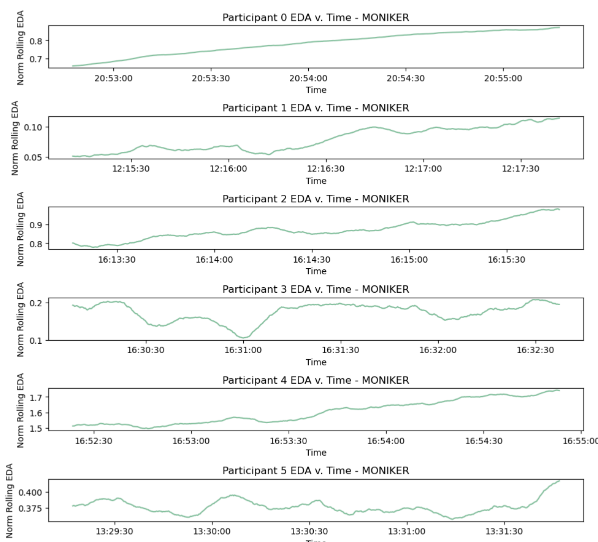


Figure 7: EDA Mean & Median

We dove into looking at the time series by category and were seeing similar results to those seen in the bar distribution above - nothing promising in terms of an overall depression or excitement of EDA for a given stimuli. This can be seen in Figure 8 which plots the time series representing normalized rolling EDA over time for each participant for a given stimuli. Very clear from both plots is the fact that each participant seems to respond differently to different images, which is seen in the ups and downs in the rolling EDA within each stimuli time series and in the different magnitudes of the bars in Figure 7. To control for different responses and feelings towards different stimuli (be it the pictures or the names or the audio) would need to be overcome by having so many recorded instances that a general trend begins to emerge as opposed to looking at only six participants.


Participant 0 EDA v. Time - CONTROL
Participant 1 EDA v. Time - CONTROL
Participant 2 EDA v. Time - CONTROL
Participant 3 EDA v. Time - CONTROL
Participant 4 EDA v. Time - CONTROL
Participant 5 EDA v. Time - CONTROL
Norm Rolling EDA
Time


Participant 0 EDA v. Time - MONIKER
Participant 1 EDA v. Time - MONIKER
Participant 2 EDA v. Time - MONIKER
Participant 3 EDA v. Time - MONIKER
Participant 4 EDA v. Time - MONIKER
Participant 5 EDA v. Time - MONIKER
Norm Rolling EDA
Time

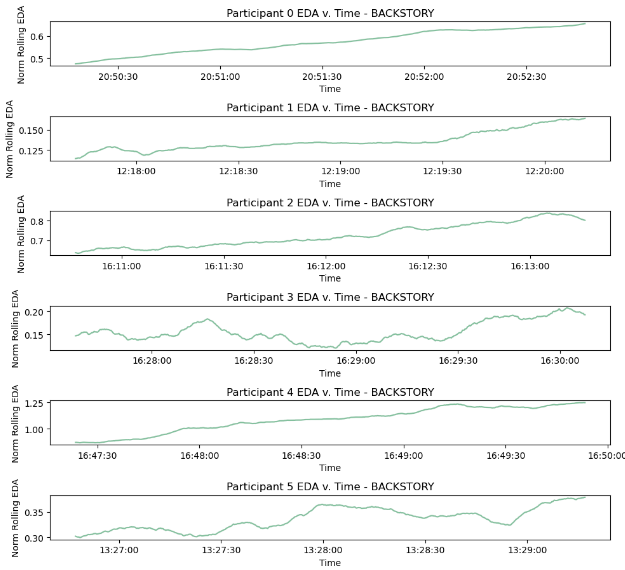


Figure 8: EDA Time Series by Stimuli Group

It was with the visualizations in Figure 8 above that we started to realize a trend that was concerning. When put in the order that the stimuli were presented, the time series represented an ever increasing EDA trend which can be seen clearly in Figure 9 below. Apart from user 4, who experienced the same experimental setup as user 3 apart from a brief interruption, the Embrace seems to take a while to ‘warm up’ which leads to a gradual increase in a user’s EDA reading over time the device is worn. The other possible interpretation of this result is that since each block is made up of variations on the same 15 images, the user’s ‘stimulation’ increases as they revisit images that they have become familiar with over time i.e. excitement increases with familiarity. The latter interpretation would need to be explored with repetition of the experiment allowing for different images through each category of stimuli or experimenting with a set-up where the user wears the sensor for a period of time prior to the introduction of the stimuli, completing a controlled activity (e.g. listening to music or meditating). Either way, the results show some response but the causal relationship is entirely inconclusive given the experimental design and number of users queried.

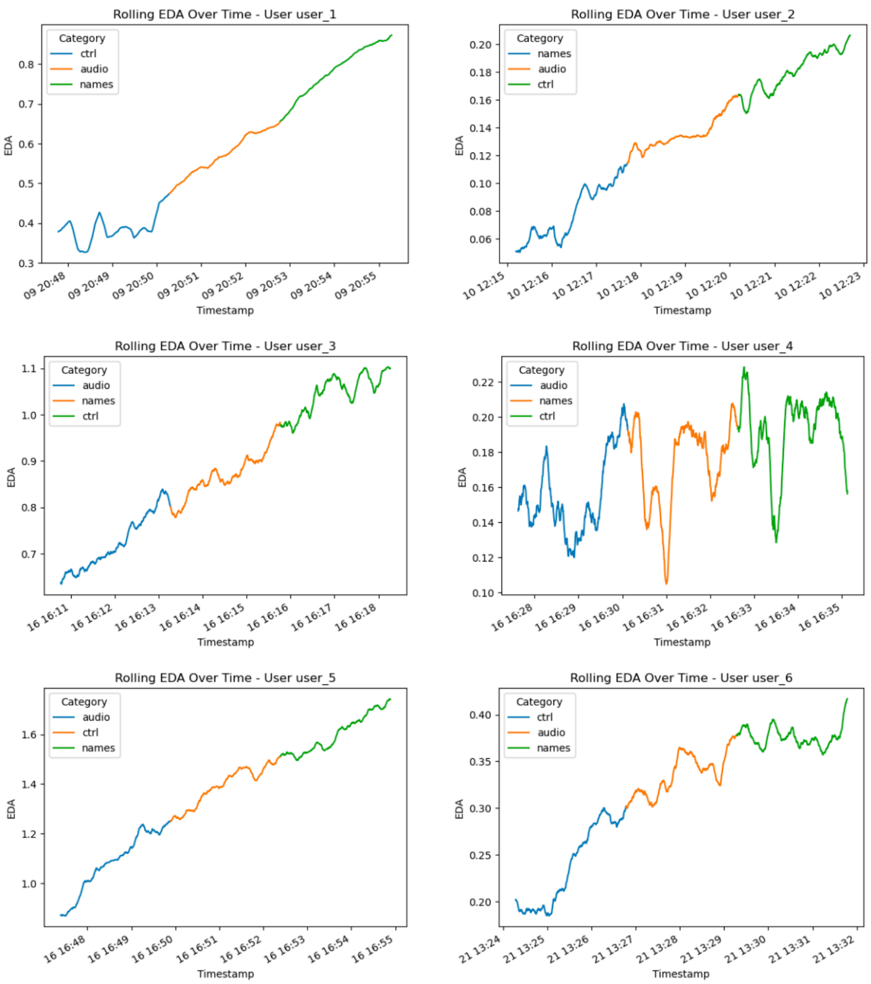
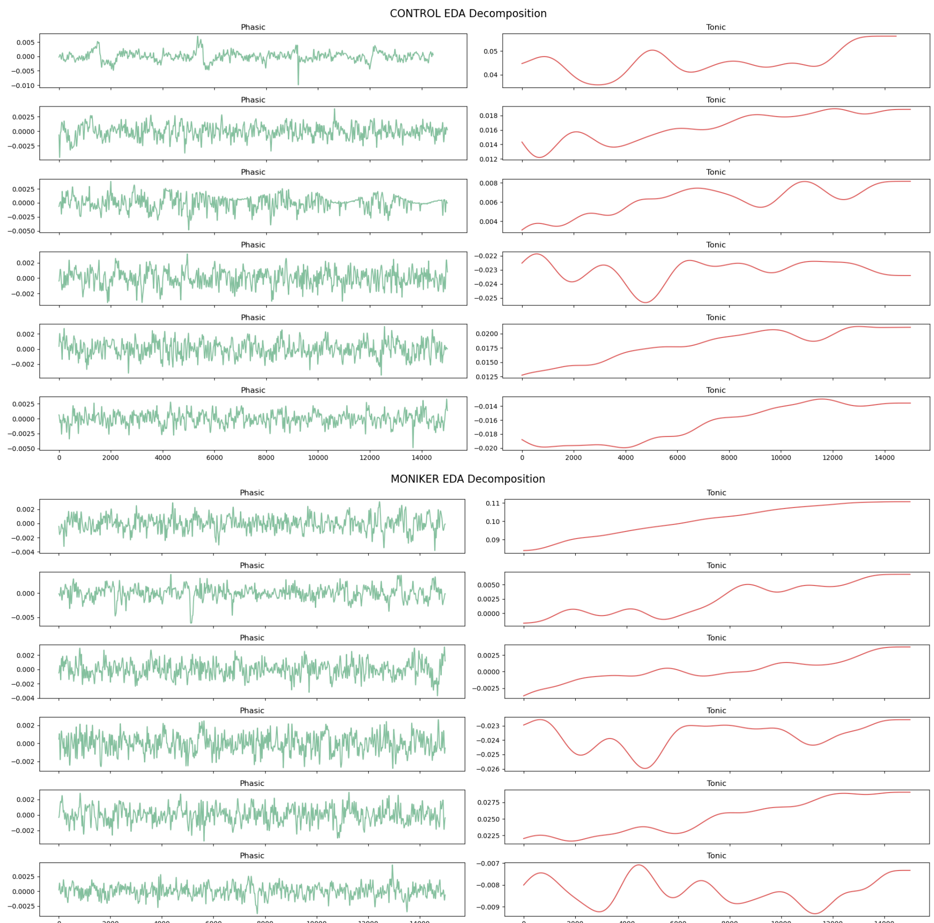


Figure 9: Rolling EDA over time

Finally, we rounded out our EDA analysis by conducting a resampling and decomposition. Resampling of the EDA data was needed due to the low sampling frequency of the EmbracePlus. By resampling, we were able to leverage the scipy.signal library and take a look at the decompositions to see if there was an overall trend in the tonic component of the signal amongst participants, which might indicate higher levels of arousal associated with a particular stimulus type. The results can be seen in Figure 10 below, which shows similar findings to that seen in the general time series plots of the rolling normalized values. Our sample size is too small and the variance due each participants response to a stimulus group or specific image is too great to truly derive meaning from these results. In the future, this work would need to be completed with a much larger sample size to effectively tease out trends in behavior or response related to a particular grouping of stimuli.



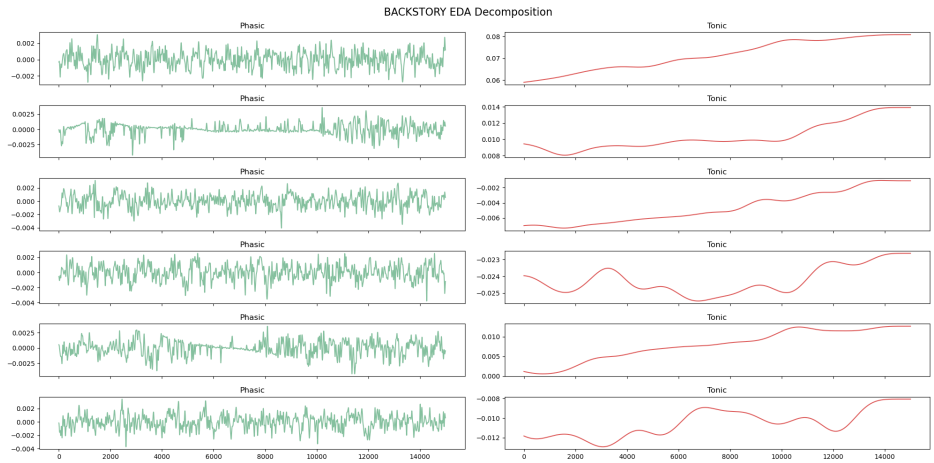


Figure 10: EDA Signal Decomposition by Stimuli Group

Discussion

The biosensory analysis yielded several promising findings and insights, albeit with some challenges. There does appear to be some EDA response and artifacts within the analysis of saccades for the gaze tracking data, however the results are incredibly preliminary and suggest the need for a more controlled experiment run with many more users and highly controlled environs.

From the EDA data, we can see each user has a different response to the stimuli. This is likely due to differences in resonance of each image with each individual, in addition to interruptions during the sampling. This is promising and could potentially be handled with future work to normalize metrics by subsampling by each individual image, however this would have taken significant set up and analysis time that was not available to our group.

The gaze data from this study opens up several fascinating areas for discussion. Firstly, it challenges pre-existing notions about how multimedia elements, specifically the combination of images, text, and audio, impact visual attention. Interestingly, the length of fixations remained relatively consistent across participants, with only one exception. This uniformity challenges the assumption that images accompanied by text and audio would inherently lead to shorter fixation durations compared to standard slides. Instead, our data implies that the presence of text does not significantly alter the duration of visual attention. Additionally, the presence of an outlier in the data highlights the importance of individual differences in information processing, suggesting a potential for more personalized content delivery based on each person's unique processing style. Secondly, the finding that Moniker and Backstory categories showed an increase in saccades suggests that the type of content can influence eye movement patterns. This opens up a discussion on how different content types, like narratives or personal stories (as in the Backstory category), might prompt viewers to explore visual information differently compared to more standard or controlled content. Understanding that certain content types do not drastically change saccade patterns could influence how information is laid out visually to optimize viewer engagement and information intake.The results could spark further research on cognitive load and attention. Specifically, it would be interesting to explore whether the increase in saccades in certain content types is due to increased cognitive engagement or simply a response to the layout and nature of the content.

One participant notably recalled the dialogue from an anthropomorphized image after 3 weeks, underscoring that the backstory element aids in memory retention. This suggests a strong correlation between the engaging nature of the backstory and the cognitive impact it has on memory. The fact that participants remembered specific dialogues points to the effectiveness of narrative elements in enhancing memory retention. However, it raises an intriguing question about whether the memory retention was a result of anthropomorphism, the humor in the dialogue, or a combination of both. Regardless, the use of anthropomorphized stimuli did appear to help our participant recall this obscure detail from our stimulus video. While anecdotal, this suggests that at the very least the motivations underpinning the research was valid, and that anthropomorphization may contribute to greater ‘attention’. The use of different categories of images to elicit varied biosensory responses appears to have been successful, particularly in terms of understanding the impact of different visual stimuli on memory retention and emotional engagement.

In the future, the experimental design could be optimized by having separate groups for each image category. This would help normalize the dataset, ensuring that the results are not influenced by the order of image presentation. Such a design would allow for a more controlled examination of the impact of each category on biosensory responses. Incorporating EEG (electroencephalogram) equipment like Unicorn, OpenBCI, or NeuroSky in future studies could provide profound insights. EEG would potentially allow for a deeper understanding of how the brain processes and interprets the data, offering a more nuanced view of the cognitive and emotional responses to different visual stimuli. This could be particularly revealing in discerning the neurological impacts of anthropomorphism, humor, and narrative elements in visual media.

References

1. Chiong, Cynthia, and Judy S. DeLoache. “Learning the ABCs: What Kinds of Picture Books Facilitate Young Children’s Learning?” *Journal of Early Childhood Literacy*, vol. 13, no. 2, June 2013, pp. 225–41. *SAGE Journals*, https://doi.org/10.1177/1468798411430091.
2. Donath, Judith S. “Anthropomorphic Visualization: Depicting Participants in Online Spaces Using the Human Form.” *MIT Media Lab*, https://www.media.mit.edu/publications/anthropomorphic-visualization-depicting-participants-in-online-spaces-using-the-human-form/. Accessed 11 Oct. 2023.
3. Li, Yiyi and Ying Xie. “Is a Picture Worth a Thousand Words? An Empirical Study of Image Content and Social Media Engagement.” *Journal of Marketing Research,* vol. 57, no. 1, 2020, pp. 1-19. *Sage Journals*, https://journals.sagepub.com/doi/10.1177/0022243719881113. Accessed 11 Oct. 2023.
4. Larsen, Nicole E., et al. “Do Storybooks with Anthropomorphized Animal Characters Promote Prosocial Behaviors in Young Children?” *Developmental Science*, vol. 21, no. 3, 2018, pp. e12590. *Wiley Online Library*, https://doi.org/10.1111/desc.12590.
5. Urquiza-Haas, Esmeralda G., and Kurt Kotrschal. “The Mind behind Anthropomorphic Thinking: Attribution of Mental States to Other Species.” *Animal Behaviour*, vol. 109, Nov. 2015, pp. 167–76. *ScienceDirect*, https://doi.org/10.1016/j.anbehav.2015.08.011.
6. Waytz, Adam, Epley, Nicholas, and John T. Cacioppo. “Social Cognition Unbound: Insights Into Anthropomorphism and Dehumanization.” *Current Directions in Psychological Science*, vol. 19, no. 1, 2020, pp. 58-62. *Sage Journals*, https://doi.org/10.1177/0963721409359302
7. Conrad, Megan, Marcovitch, Stuart, and Janet J. Boseovski, “The friendly fossa: The effect of anthropomorphic language on learning about unfamiliar animals through both

storybooks and live animal experiences." *Journal of Experimental Child Psychology.* vol. 201, no. 1, January 2021, pp. 1-20. *ScienceDirect,* https://doi.org/10.1016/j.jecp.2020.104985

8. Conrad, Megan. "(Un)Real Animals: Anthropomorphism and Early Learning About Animals." *Child Development Perspectives*. vol. 10, no. 1, 2016, pp. 10 - 14. *Society for Research in Child Development,* https://doi.org/10.1111/cdep.12153
9. Baker, Lewis J., Hymel, Alicia M., and Daniel T. Levin. "Anthropomorphism and Intentionality Improve Memory for Events." *Discourse Processes,* vol. 55, issue 3, September 2016, pp. 241-255, *Taylor & Francis Online,* https://doi.org/10.1080/0163853X.2016.1223517
10. Kallery, Maria, Psillos, Dimitris. "Anthropomorphism and Animism in Early Years Science: Why Teachers Use Them, how They Conceptualise Them and What Are Their Views on Their Use." *Research in Science Education* vol. 34, 2004, pp. 291–311. https://doi.org/10.1023/B:RISE.0000044613.64634.03
11. Brossard Stoos, Kari A., Haftel, Madeline. "Using Anthropomorphism and Fictional Story Development to Enhance Student Learning." *J Microbiol Biol Educ.* vol. 18 n.1.19. Apr 21, 2017. doi: 10.1128/jmbe.v18i1.1197. PMID: 28912934;
12. Allgood, Beth, and Dehara Weeraman. "Well-Being for All Beings." *OneNature*, 21 Apr. 2023, onenatureinstitute.org/. https://onenatureinstitute.org/wp-content/uploads/2023/05/Screenshot-2023-05-03-at-3.42.26-PM-1080x576.png
13. *Watercolor Monkey*. *One Nature Institute*, https://onenatureinstitute.org/wp-content/uploads/2023/05/Screenshot-2023-05-03-at-3.42.26-PM-1080x576.png.

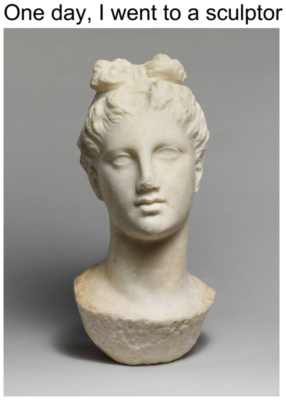
Appendix

Appendix A - Stimulus Images, Monikers, and Stories

1. **Pet Rock**
   1.1 Name: Hi! My name is Rocky.
   1.2 Story: Hi! My name is Rocky. It took me a lot of rolling to get here. I'm excited to meet you!
   
2. **Dog**
   2.1 Name: Woof! I'm Jasper.
   2.2 Story: Woof! I'm Jasper. My owners left for work, and I'm not sure when they're coming back. I'd love for someone to play with me.
   
3. **African Mask**
   3.1 Name: Hello, I'm Matthew.
   3.2 Story: Hello, I'm Matthew. I was carved by the most skilled artisan that I know. A

piece like me can only come from meticulous craftsmanship and care.

4. **Marble Head**
   4.1 Name: Good Day. I am Diogenes.
   4.2 Story: Good Day. I am Diogenes. I always dreamt of what immortality looked like. One day, I went to a sculptor to make my dream a reality.



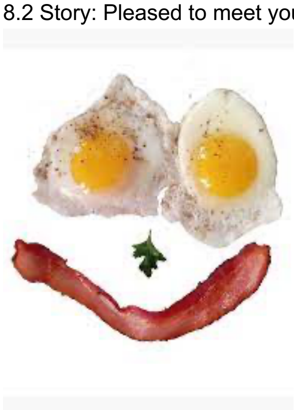
5. Cubism Face
   5.1 Name: Greetings - I am Sarah.
   5.2 Story: Greetings - I am Sarah. Grief, curiosity, and love lost etch the colors on my canvas.

6. **Architecture**
   6.1Name: I. Am. Frederick.
   6.2 Story: I. Am. Frederick. They built me to house a family. They built us to house a community.

7. **Vase**
   7.1 Name: Helloooo dear, my name is Wilma!
   7.2 Story: Helloooo dear, my name is Wilma! He said that I’m meant to be celery, but I’m so much more than that darling.

8. **Eggs and Bacon**
   8.1 Name: Pleased to meet you, I’m Eggsy.

8.2 Story: Pleased to meet you, I’m Eggsy. I’m here to start your day right!



9. Graffiti

9.1 Name: Daryll, at your service.

9.2 Story: Daryll, at your service. You like what you see? I’m pretty eye-catching aren’t I?

10. Candy

10.1 Name: Hiya! My name is Phinneas.

10.2 Story: Hiya! My name is Phinneas. I might not be the healthiest snack, but hey the kids love me.

11. Cat

11.1 Name: Meow! They call me Toby.

11.2 Story: Meow! They call me Toby. I can tell by looking at you that you want to pet me. Well, you can't.

12. Car

12.1 Name: Hey there! I'm Felicia.

12.2 Story: Hey there! I'm Felicia. You ever felt your hair blow in the wind while riding down the coast? No? Hop in baby let's go for a ride.

13. Socket

13.1 Name: What's Up! My name is Wally.

13.2 Story: What's Up! My name is Wally. I usually am a pretty chill dude, but make sure you don't overload me or I might blow a fuse!

14. Tree

14.1 Name: Howdy - I'm Karl.

14.2 Story: Howdy - I'm Karl. What are you doing in my woods? These are not your woods. You do not need to be here.

15. Coffee

15.1 Name: Hiiiiiii! I'm Lenny!

15.2 Story: Hiiiiiii! I'm Lenny! Drink me! I'll wake you up! Drink me drink me!

Appendix B - Images of our Experimental Procedure

1. Tried using Neurosky, had bluetooth configuration issues with windows laptop

2. Participant wearing Empatica and Pupil Invisible. We had 2 QR codes to mark the screen for gaze analysis